The normative development of child and adolescent problem behavior. [IF 3.2]

Bongers, I.L.; Koot, H.M.; van der Ende, J.; Verhulst, F.C.

published in
Journal of Abnormal Psychology
2003

DOI (link to publisher)
10.1037/0021-843X.112.2.179

document version
Publisher's PDF, also known as Version of record

Journal of Abnormal Psychology
2003, Vol. 112, No. 2, 179–192

0021-843X/03/$12.00 DOI: 10.1037/0021-843X.112.2.179

The Normative Development of Child and Adolescent Problem Behavior

Ilja L. Bongers
Erasmus MC-Sophia Children's Hospital

Hans M. Koot
Erasmus MC-Sophia Children's Hospital
and Free University Amsterdam

Jan van der Ende and Frank C. Verhulst
Erasmus MC-Sophia Children's Hospital

The aim of this study was to identify normative developmental trajectories of parent-reported problems assessed with the Child Behavior Checklist (CBCL; T. M. Achenbach, 1991) in a representative sample of 2,076 children aged 4 to 18 years from the general population. The trajectories were determined by multilevel growth curve analyses on the CBCL syndromes in a longitudinal multiple birth-cohort sample that was assessed 5 times with 2-year intervals. Most syndromes showed a linear increase or decrease with age or a curvilinear trajectory, except for thought problems. Trajectories for most syndromes differed for boys versus girls, except those for withdrawn, social problems, and thought problems. These normative developmental trajectories provide information against which developmental deviance in childhood and adolescence can be detected.

Little is known about the normative development of behavioral and emotional problems of children and adolescents in the general population. Although several studies and review articles provide suggestive evidence of age-related changes in problem behavior, this evidence is mainly derived from cross-sectional studies or single birth-cohort longitudinal studies (Birmaher et al., 1996; Campbell, 1995; Cicchetti & Toth, 1998; Fergusson, 1998). However, for the determination of developmental trajectories, multicohort longitudinal studies are necessary because they enable the researcher to disentangle age, cohort, and period effects by showing whether the same changes with age are observed in different cohorts studied in different time periods (Farrington, 1991; Kraemer, Yesavage, Taylor, & Kupfer, 2000; Loeber & Farrington, 1995; Willett, Singer, & Martin, 1998).

Only two studies have investigated the normative development of psychopathology in the general population (Keiley, Bates, Dodge, & Pettit, 2000; Stanger, Achenbach & Verhulst, 1997). The study by Stanger et al. (1997) is the only study using a multicohort longitudinal sample, including 1,139 children ages 4–18 years. The authors studied the normative developmental trajectories of delinquent and aggressive behavior using an accelerated longitudinal design with five repeated measurements at 2-year intervals with data that partially overlapped data used in the present study. This design matches two cohorts with at least two overlaps in measurement moments. The overlapping cohorts were matched by level of problem behavior, gender, socioeconomic status (SES), and age. According to this study, the scores for both the Aggressive and the Delinquent Behavior measures declined from ages 4 to 10 years. After about the age of 10 years, scores for Aggressive Behavior continued to decline, but scores for Delinquent Behavior increased until age 17. On both aggressive and delinquent behaviors, boys were scored higher than girls. A disadvantage was that Stanger et al.'s design did not control for the effects of an interaction between age and cohort. This effect can only be controlled in a nonmatched multicohort longitudinal study (Raudenbush & Chan, 1992).

The study by Keiley et al. (2000) used a single birth-cohort longitudinal design. More than 400 children aged 5–12 years from the general population were followed from kindergarten through seventh grade and were assessed every year (eight measurements). Using multilevel data analytic techniques, the study determined the developmental trajectories of internalizing and externalizing problems of these children reported by the mother and the teacher. For internalizing behaviors reported by both mothers and teachers, no effect of gender and time was found, which suggests a stable trajectory for internalizing behavior throughout the measurement period. However, mothers and teachers differed in their reports of externalizing behavior. Teachers reported significantly more externalizing behaviors in boys than in girls and a decline over time that was faster for boys than for girls. A similar declining trajectory was found for mother-reported externalizing behavior, but it is surprising that no effect of gender was found. However, the use of a single cohort implies that period and cohort effects could not be controlled for in this study.

Ilja L. Bongers, Jan van der Ende, and Frank C. Verhulst, Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands; Hans M. Koot, Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia Children's Hospital, and Department of Developmental Psychology, Free University Amsterdam, Amsterdam, the Netherlands.

This research project was supported by Sophia Scientific Foundation Grant 302.

Correspondence concerning this article should be addressed to Ilja L. Bongers, Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia Children's Hospital, Dr. Molewaterplein 60, 3015 GJ Rotterdam, the Netherlands. E-mail: i.l.bongers@erasmusmc.nl

Neither study accounted for age, period, or cohort effects in the way a multicohort longitudinal study can. The present study addresses the normative development of psychopathology using multiple cohorts and taking advantage of multilevel growth curve analysis, a data-analytic method that was specifically developed to describe time-related changes while accounting for cohort and period effects.

This study used parent reports of children's problems on the Child Behavior Checklist (CBCL; Achenbach, 1991; Verhulst, Van der Ende, & Koot, 1996) obtained at five time points with 2-year intervals. The analysis aimed to estimate the normative developmental trajectories for all CBCL scores, including Total Problems, Internalizing, Externalizing, and eight small-band scale scores (i.e., Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior). The normative developmental trajectories represent the changes in levels of problem behavior from age 4 to 18 years. We examined the initial status and age-related change of the behavior as well as gender differences in these parameters.

On the basis of the limited evidence available, a number of hypotheses on age- and gender-related changes in problem behavior may be proposed. Because we aimed to describe the normative development of problem behavior, we reviewed only studies including samples drawn from the general population. We organized our hypotheses according to the problem scales identifiable in the CBCL (Achenbach, 1991; Verhulst et al., 1996), because the data in this study were obtained with this instrument.

Internalizing Problems

Internalizing problems include anxiety, depression (Bernstein, Borchardt, & Perwien, 1996; Birmaher et al., 1996), somatic complaints (Egger, Costello, Erkanli, & Angold, 1999; Taylor, Szatmari, Boyle, & Offord, 1996), and withdrawn behavior.

Symptoms of anxiety are quite common in childhood and adolescence, but their type and content vary with age (Craske, 1997). Anxiety problems change from separation anxiety in early childhood to social phobia or generalized anxiety in adolescence. Specific phobia has been described in children of all ages. However, for the total number of anxiety symptoms, no consistent gender or age differences have emerged (Bell-Dolan, Last, & Strauss, 1990; Bernstein et al., 1996).

For symptoms of depression, consistent age and gender differences have been found (e.g., Angold & Rutter, 1992). Prepubertal boys and girls show equal levels of depressive problems, but around midpuberty (Tanner Stage III) girls begin to exhibit more depressive problems, a trend that continues into adulthood (Angold, Costello, & Worthman, 1998; Birmaher et al., 1996; Cohen et al., 1993; Fleming & Offord, 1990; Laitinen-Krispijn, Van der Ende, & Verhulst, 1999).

In the CBCL, the anxious and depressed behaviors are subsumed under one construct (i.e., the Anxious/Depressed scale). Whereas findings suggest an increase of symptoms of depression over age, especially in girls, anxiety is expected to be stable over ages. Because anxiety is more normative than depression, the normative developmental trajectory of the anxious/depressed syndrome is expected to be similar to the trajectory of anxiety. Thus, we expected a small increase in scores on the Anxious/Depressed scale over age for girls and a stable (i.e., neither increasing nor decreasing) normative developmental trajectory for boys.

Another syndrome included in the CBCL Internalizing scale is the somatic complaints syndrome. Children and adolescents do not differ in level of self-reported somatic complaints (Taylor et al., 1996). Girls report more somatic complaints than do boys, and this difference continues into adulthood (Egger et al., 1999; Taylor et al., 1996). Thus, the developmental trajectory of the somatic complaints syndrome is expected to stay at the same level across childhood and adolescence and is expected to have higher levels for girls than for boys.

Finally, the CBCL Internalizing scale includes the withdrawn syndrome. Scores on this syndrome are expected to show an increase from childhood to adolescence. Young adolescents usually acquire more independence from their parents, and peers become more important as a reference group (Alsaker, 1996). This increases the emotional distance between adolescents and their parents. Pubertal maturation is related to greater emotional autonomy and less closeness to parents (Alsaker, 1996). From cross-sectional studies (Achenbach, 1991; Verhulst et al., 1996) it is known that parents of older children report more withdrawn behavior than do parents of younger children and that parents report more withdrawn behaviors for girls than for boys (Achenbach, 1991). These findings suggest a normative developmental trajectory for the Withdrawn scale with scores that are increasing with age for both girls and boys, with higher overall levels for girls than for boys.

Externalizing Problems

The CBCL broadband Externalizing scale encompasses the syndromes aggressive behavior and delinquent behavior. The aggressive behavior syndrome includes behaviors such as bragging, teasing, fighting, and attacking, which may indicate both aggression and opposition (Frick et al., 1993). Young children are still developing their communication skills and often rely on aggressive and oppositional behavior to control their environment (Tremblay, 2000). The type of aggressive behavior that both boys and girls show transforms during development, with decreasing levels of physical aggression (Cairns, Cairns, Neckerman, Ferguson, & Gariépy, 1989). Most studies indicate that boys show more physical and verbal aggression than girls do (Cairns & Cairns, 1984; Silverthorn & Frick, 1999). For both boys and girls, we expected a declining normative developmental trajectory of scores on the Aggressive Behavior scale, with higher levels for boys versus girls.

The delinquent behavior syndrome includes behaviors such as stealing and fire setting but also lying and cheating. These acts are covert and may be both destructive and nondestructive (Frick et al., 1993). Recent studies (Loeber et al., 1993; Moffitt, Caspi, Dickson, Silva, & Stanton, 1996; Nagin & Tremblay, 1999) show that the majority (60–70%) of boys do not commit any delinquent or antisocial acts during childhood and adolescence. For girls this group is larger (i.e., nearly 90% of girls never commit any delinquent or antisocial act; Moffitt & Caspi, 2001). It has been suggested that there is a small group of boys (10%) and girls (1%) who show persistent delinquent and antisocial behaviors throughout childhood and adolescence. They follow a so-called life-course-persistent trajectory (Moffitt, 1993). A larger group of boys

(26%) and girls (18%) only commit antisocial or delinquent behaviors during adolescence. They follow the so-called adolescence-limited trajectory (Moffitt, 1993). These findings suggest a normative developmental trajectory of increasing delinquent behavior starting in adolescence, with boys showing more delinquent behavior problems than girls across the whole period of childhood and adolescence.

Attention Problems

Attention problems are supposed to emerge when children are starting to attend school, because there children are faced with more complex and structured tasks. A few recent studies provide evidence for the expected normative development of attention problems. A study on a general population sample in Australia indicated minimal age differences in the number of attention deficit hyperactivity disorder symptoms in the age range from 5 to 11 years (Gomez, Harvey, Quick, Scharer, & Harris, 1999). Results from a study in a clinically referred male sample indicated that hyperactivity–impulsivity declines with increasing age, especially during late childhood and early adolescence, whereas inattention remains relatively stable over ages (Hart et al., 1995). Other studies indicated that boys show more symptoms of inattentive and overactive behaviors than do girls (Cantwell, 1996; Gaub & Carlson, 1997; Gomez et al., 1999). These findings suggest a normative trajectory with first an increase when children start attending school and thereafter a decreasing frequency of attention problems over age, with higher levels for boys than for girls.

Other Problems

Two other CBCL syndromes not included in the Internalizing and Externalizing scales are the social problems and thought problems syndromes. Because these syndromes have no direct counterparts in the general psychopathology literature that does not use the CBCL, our hypotheses have to be based on cross-sectional findings reported for these syndromes.

The social problems syndrome includes behaviors such as acting too young, getting teased, or not being liked by peers. Results from cross-sectional studies indicate that younger children show more social problems than do older ones and that boys show more problems than do girls (Achenbach, 1991; Verhulst et al., 1996). These cross-sectional findings suggest a normative developmental trajectory that decreases with age and has a higher level for boys than for girls.

The thought problems syndrome includes items referring to obsessive–compulsive behavior, seeing or hearing things that are not there, and strange behaviors. Results from cross-sectional studies indicate no age effects or gender effects on the thought problems syndrome (Achenbach, 1991; Verhulst et al., 1996), which suggests that the normative developmental trajectory for thought problems is stable during childhood and adolescence and similar for boys and girls.

Method

Sample

The data used in this study were derived from a six-wave longitudinal study of behavioral and emotional problems that began in 1983. Respondents were interviewed at 2-year intervals until 1991 and again in 1997. This study uses data from the first five waves. The original sample of 2,600 children from 13 birth cohorts aged 4 to 16 years was drawn from municipal registers that list all residents in the Dutch province of Zuid-Holland. A random sample was drawn of 100 children of each gender and age with the Dutch nationality. Two small municipalities out of a total of 86 refused to cooperate, and 75 children were untraceable. Of the 2,447 parents who could be reached, 2,076 responded and provided usable CBCL data (84.8%). For details of the initial data collection, see Verhulst, Akkerhuis, and Althaus (1985). After the first measurement (in 1983), the sample was approached again in 1985, 1987, 1989, and 1991. The sample at Time 1 included 1,016 boys and 1,060 girls (see Table 1).

Because of the age range of the CBCL, not all subjects could participate in each wave of the study. The age range was 4 to 16 years at Time 1 and Time 2 for the earlier version of the CBCL and 4 to 18 years at Time 3

Table 1
Number of Subjects by Time of Measurement

	Time 1, 1983 4–16 yrs			Time 2, 1985 6–18 yrs			Time 3, 1987 8–20 yrs			Time 4, 1989 10–22 yrs			Time 5, 1991 12–24 yrs		
Cohort	Age	B	G	Age	B	G	Age	B	G	Age	B	G	Age	B	G
1	4	81	84	6	69	64	8	69	71	10	72	76	12	71	75
2	5	78	90	7	65	73	9	65	75	11	70	73	13	69	73
3	6	78	83	8	63	71	10	65	70	12	65	71	14	68	74
4	7	78	85	9	67	72	11	65	71	13	67	74	15	66	71
5	8	89	83	10	66	62	12	77	64	14	72	69	16	67	65
6	9	81	78	11	66	72	13	60	71	15	67	73	17	62	67
7	10	78	83	12	59	63	14	59	67	16	64	69	18	60	66
8	11	78	83	13	65	66	15	66	69	17	62	72	19		
9	12	77	76	14	63	61	16	55	63	18			20		
10	13	78	83	15	52	65	17	43	51	19			21		
11	14	69	82	16	50	57	18	37	41	20			22		
12	15	75	70	17	1		19			21			23		
13	16	76	80	18			20			22			24		
Total		1,016	1,060		686	726		661	713		539	577		463	491

Note. yrs = years; B = boys; G = girls.

through Time 5 for the 1991 version of the CBCL. Of the 2,076 subjects who participated at Time 1, only 1,149 (Cohorts 1 to 7) were able to participate at all five waves; of these subjects, 68.8% completed the CBCL at all five time points (see Table 1). Of the 2,076 who participated at Time 1, data were available for 38.1% from five measurements, for 12.2% from four measurements, for 18.0% from three measurements, for 8.8% from two measurements, and for 22.9% from only one measurement. We kept all subjects in the sample who were between 4 to 18 years of age at any time point even if data were available from only one measurement.

To investigate selective attrition, we compared dropouts and remainders with respect to their Time 1 CBCL Total Problems score and SES. We divided the sample into three groups, one group with subjects who participated in five waves (remainders), one group with one or more missing waves scattered throughout the study (random dropouts), and one group of subjects who participated only once, at the first wave, and never participated again (dropouts). SES of the parents at Time 1 was scored on a six-step scale of parental occupation (Van Westerlaak, Kropman, & Collaris, 1975), with 1 indicating the lowest SES and 6 indicating the highest SES. We examined the differences in the mean SES and the CBCL Total Problems score between the dropout groups using analyses of covariance (ANCOVAs) with age as a covariate and dropout as a fixed factor, so that the effect of the dropout groups on the SES and the Total Problems score was corrected for the age of the individual. The ANCOVAs resulted in a significant effect of dropout in the mean SES, $F(2, 2064) = 10.314$, $p < .01$, and no effect of dropout in the mean Total Problems score, $F(2, 2071) = 0.065$, $p = .937$. Mean SES (adjusted) was 3.77 for the remainders, 3.53 for the random dropouts, and 3.19 for the dropouts.

Measurements

At Time 1 to Time 5, the CBCL (Achenbach, 1991) was used to obtain standardized parent reports of children's problem behaviors. Of the collected CBCLs, 87% were filled out by the mother. The CBCL is a questionnaire to be completed by parents of 4- to 18-year olds and contains 120 items covering behavioral or emotional problems that occurred during the past 6 months. The response format is 0 (*not true*) through 2 (*very true or often true*). The CBCL can be scored on the syndrome scales: Anxious/Depressed, Withdrawn, Somatic Complaints (these three scales form the Internalizing scale), Aggressive Behavior, Delinquent Behavior (these two scales form the Externalizing scale), Attention Problems, Social Problems, and Thought Problems. One can derive a Total Problems score by summing the individual item scores. The same items and syndromes are scored for boys and girls aged 4 to 18 years. The good reliability and validity of the CBCL (Achenbach, 1991) were confirmed for the Dutch version of the measure (Verhulst et al., 1985, 1996). Cronbach's alpha ranged from .92 for Total Problems score to .40 for Delinquent Behavior (average = .70). The test–retest reliability over a period of 2 weeks ranged from .91 for Total Problems score to .74 for Thought Problems (Verhulst et al., 1996). Confirmatory factor analysis of the American syndromes in a sample of 2,335 clinically referred Dutch children, aged 4 to 18 years, supported the cross-cultural generalizability of the CBCL (De Groot, Koot, & Verhulst, 1994). The 4-year stability ranged from .65 for Aggressive Behavior to .24 for Thought Problems (Verhulst, Koot, & Berden, 1990).

Statistical Analyses

The normative course of the CBCL syndromes, the two broadband groupings, and Total Problems scores were described with multilevel growth curve analysis (Bryk & Raudenbush, 1992). Multilevel models deal with the analysis of nested data. In a multiwave longitudinal sample, the repeated observations are nested within the individuals. Each wave contributes to a naturally formed subset of observations that are nested within individuals. The multilevel model has two levels: one level for the repeated measures (Level 1 or between subjects), and one level for the individuals (Level 2 or within subject). Level 1, the level for the repeated measures, describes the between-subjects variation with the use of the following parameters: intercept, gender, and age. The values obtained for the parameters included in Level 1 describe the normative developmental trajectories. Level 2, the level for the individuals, describes the characteristics of the individuals who participate in the study (i.e., the within-subject variation). In multilevel models, each individual is allowed his or her own growth curve or growth trajectory—that is, the individual growth parameters (intercept and slope) may vary across individuals—and these parameters are modeled in the Level 2 part of the multilevel growth curve model.

In this study, the within-subject variation is described by the intercept of the scale, the gender, and the age of the subject. An overall growth curve is estimated for the total sample that is the average of all individual growth curves. The parameters that describe the overall growth curve are estimated in the Level 1 part of the multilevel growth curve model. In that way, the multilevel model allows for estimation of the mean growth trajectory (for the total sample) as well as the estimation of individual variation around this mean. The multilevel growth trajectory at each level consists of the growth parameters' intercept and slope. The intercept represents the initial status of the problem behavior at age 4 on both the between-subjects and the within-subject level. The slope describes the average rate of change in problem behavior across ages for each individual on both levels (Boyle & Willms, 2001; Bryk & Raudenbush, 1992).

The multilevel growth curve models were estimated with (restricted) maximum likelihood estimation and an unstructured (co)variance matrix according to the SAS PROC MIXED procedure (Littell, Milliken, Stroup, & Wolfinger, 1996). The unstructured (co)variance matrix gave a better fit for the (co)variance structure than the compound symmetry model and the first-order autoregressive model. In the unstructured (co)variance matrix the (co)variances are independent of each other and uncorrelated. We used the raw scale scores to estimate the multilevel growth curves, following the same procedure for each syndrome. First we fitted the baseline models, which consist of only the intercept at Level 1 and Level 2. After estimating the baseline model, we tested which of the nested Level 1 models gave the best fit using the maximum likelihood chi-square difference test. We created seven different nested models with the following parameters: gender, age, age squared, and Gender × Age. All nested models had an intercept. After deciding which model described the (Level 1) mean growth curve best, we built the Level 2 of the multilevel growth curve model in the same way. We used the restricted maximum likelihood chi-square instead of the normal maximum likelihood chi-square because the former is an unbiased estimator of the covariance matrix (Longford, 1993). In both levels we tested the significance of the nested models instead of the significance of the parameters. Therefore, it is possible that parameters included in the final models are not significant. However, the model with nonsignificant parameters showed a better model fit than did models without these parameters.

The multilevel analysis deals with the missing data by the expectation–maximization (EM) algorithm. The EM algorithm is an iterative procedure, with each iteration consisting of two parts: the E step (expectation), in which the conditional expectations of the functions of the sufficient statistics for the missing data are calculated, and the M step (maximization), in which the complete likelihood, with the functions of the missing data replaced by their conditional expectations, is maximized. The conditioning in the E step is on the incomplete (available) data and the current estimates of the parameters. The estimates are updated in the subsequent M step. This process of E steps and M steps continues until the estimates of the parameters stop changing to a meaningful extent.

The age parameter was centered at age 4 before being entered into the model. This causes the estimates of the parameters in the model to be tested for significance at age 4. Gender was coded as a dummy variable, with boys = 1 and girls = 0. Because of this dummy variable, the value for the intercept corresponds to the raw score on age 4 for girls. The intercept plus the gender effect is the average initial status of the growth trajectory. A significant gender effect indicates that there is a significant difference

between boys and girls in the level of problem behavior at age 4. The following equation represents a full model with all the possible parameters on both levels (γ parameters represent the between-subjects level parameters, and μ parameters represent the within-subject level parameters):

$$\text{CBCL scale score} = \gamma_{00} + \gamma_{10}*\text{gender} + \gamma_{20}*\text{age} + \gamma_{30}*\text{Gender} \times \text{Age} + \gamma_{40}*\text{age}^2 + \mu_{0j} + \mu_{1j}*\text{gender} + \mu_{2j}*\text{age} + \mu_{3j}*\text{Gender} \times \text{Age} + \mu_{4j}*\text{age}^2 + \varepsilon_{ij}. \quad (1)$$

To test the model fit of the models, we calculated the root-mean-square error of approximation (RMSEA; Steiger, 1998). An RMSEA value of .05 or smaller suggests that the absolute magnitude of the discrepancies between the models and the data is small (Hu & Bentler, 1999). Values in the range of .05 to .08 indicate a fair fit, and values above .10 indicate a poor fit. We consider values in the range of .09 to .10 to indicate mediocre fit (MacCallum, Browne, & Sugawara, 1996).

Results

We first calculated the means and standard deviations of all CBCL syndromes and the Total Problems score, separately for boys and girls, and for four age groups. As shown in Table 2, the means vary by the number of items in each scale and show increases and decreases by age. Next, we tested multilevel growth curve models for all syndromes and for the Total Problems score. Table 3 shows for each syndrome the estimated parameters of the

Table 2
Means and Standard Deviations of the Raw Child Behavior Checklist Syndrome Scores by Age Category

	Boys				Girls			
Syndrome	4–7 yrs	8–11 yrs	12–14 yrs	15–18 yrs	4–7 yrs	8–11 yrs	12–14 yrs	15–18 yrs
Anxious/Depressed								
14 items								
M	1.81	2.63	2.17	1.97	1.93	2.47	2.50	2.72
SD	2.48	3.18	2.84	2.84	2.45	2.99	3.19	3.57
Somatic Complaints								
9 items								
M	0.61	0.71	0.75	0.76	0.67	0.81	1.00	1.28
SD	1.14	1.38	1.33	1.36	1.13	1.27	1.59	1.93
Withdrawn								
9 items								
M	1.58	1.81	1.93	2.12	1.78	1.94	2.04	2.27
SD	1.81	1.99	2.10	2.29	1.83	2.11	2.08	2.35
Internalizing								
32 items								
M	3.96	5.09	4.78	4.78	4.35	5.15	5.45	6.11
SD	4.13	5.09	4.88	5.15	4.16	4.98	5.40	6.19
Aggressive Behavior								
20 items								
M	8.50	7.12	5.76	4.52	6.28	5.05	4.47	3.81
SD	6.28	6.08	5.42	4.98	5.24	5.04	4.70	4.41
Delinquent Behavior								
13 items								
M	1.47	1.28	1.25	1.34	1.03	0.84	0.91	1.01
SD	1.69	1.68	1.94	2.10	1.24	1.31	1.58	1.72
Externalizing								
33 items								
M	9.97	8.40	7.01	5.86	7.30	5.90	5.38	4.82
SD	7.41	7.21	6.86	6.52	6.08	5.89	5.85	5.64
Attention Problems								
11 items								
M	3.16	3.78	3.57	3.12	2.35	2.74	2.72	2.59
SD	2.70	3.37	3.20	3.18	2.47	2.88	2.80	2.91
Social Problems								
8 items								
M	1.30	1.51	1.40	1.04	1.15	1.44	1.25	0.99
SD	1.75	2.06	2.00	1.72	1.59	1.98	1.80	1.67
Thought Problems								
7 items								
M	0.36	0.29	0.23	0.25	0.26	0.24	0.25	0.28
SD	0.83	0.73	0.66	0.73	0.64	0.63	0.77	0.79
Total Problems								
118 items								
M	22.84	21.90	18.99	16.52	19.58	18.60	17.21	16.54
SD	15.08	16.52	15.47	15.14	13.96	14.60	14.69	15.31

Note. yrs = years.

Table 3
Model Fit and Parameter Estimates for the Final Models at the Between-Subjects and Within-Subject Level

	Model fit					Between-subjects level		Within-subject level	
Syndrome	Final χ^2 (*df*)	Baseline χ^2 (*df*)	$\Delta\chi^2$	Δdf	RMSEA (CI)	Parameter	Estimates (*SE*)	Parameter (co)variances	Estimates (*SE*)
Anxious/ Depressed	32,604.3 (2065)	32,875.3 (2074)	271.0***	9	0.08 (0.07–0.09)	Intercept	1.73*** (0.14)	Intercept	2.16*** (0.51)
						Gender	0.26 (0.18)	Age	0.022 (0.018)
						Age	0.19*** (0.031)	Age^2	0.000
						Gender × Age	−0.076*** (0.019)	Intercept × Age	0.51*** (0.096)
						Age^2	−0.0093*** (0.0019)	Intercept × Age^2	−0.055*** (0.0065)
								Age × Age^2	0.0016 (0.00062)
Somatic Complaints	23,498.4 (2066)	23,806.9 (2074)	308.5***	8	0.07 (0.06–0.08)	Intercept	0.57*** (0.054)	Intercept	0.000
						Gender	0.054 (0.078)	Age	0.030** (0.011)
						Age	0.054*** (0.0064)	Age^2	0.00024*** (0.000059)
						Gender × Age	−0.042*** (0.0092)	Intercept × Age	0.057** (0.021)
								Intercept × Age^2	−0.0026 (0.0020)
								Age × Age^2	−0.0025** (0.00080)
Withdrawn	27,541.4 (2068)	27,754.6 (2074)	213.2***	6	0.08 (0.07–0.08)	Intercept	1.70*** (0.063)	Intercept	1.38*** (0.25)
						Age	0.037*** (0.0066)	Age	0.015 (0.0093)
								Age^2	0.000
								Intercept × Age	0.19*** (0.047)
								Intercept × Age^2	−0.023*** (0.0031)
								Age × Age^2	0.00051 (0.00031)
Internalizing	39,576.8 (2065)	39,876.2 (2074)	299.4***	9	0.09 (0.08–0.10)	Intercept	4.10*** (0.23)	Intercept	5.17*** (1.22)
						Gender	0.0056 (0.29)	Age	0.000
						Age	0.27*** (0.049)	Age^2	0.000
						Gender × Age	−0.099** (0.031)	Intercept × Age	1.71*** (0.19)
						Age^2	−0.0098** (0.0031)	Intercept × Age^2	−0.16*** (0.016)
								Age × Age^2	0.0064*** (0.00061)
Aggressive Behavior	38,955.9 (2065)	39,856.0 (2074)	900.1***	9	0.09 (0.08–0.10)	Intercept	6.86*** (0.31)	Intercept	33.13*** (2.95)
						Gender	2.98*** (0.37)	Age	0.47** (0.13)
						Age	−0.31*** (0.055)	Age^2	0.000029 (0.00053)
						Gender × Age	−0.19*** (0.032)	Intercept × Age	−1.72** (0.54)
						Age^2	0.0059 (0.0031)	Intercept × Age^2	−0.040 (0.033)
								Age × Age^2	−0.0097 (0.0084)
Delinquent Behavior	25,073.8 (2066)	25,341.1 (2074)	267.3***	8	0.07 (0.07–0.08)	Intercept	1.17*** (0.072)	Intercept	0.69*** (0.15)
						Gender	0.37*** (0.061)	Age	0.020** (0.0071)

Table 3 (*continued*)

	Model fit					Between-subjects level		Within-subject level	
Syndrome	Final χ^2 (*df*)	Baseline χ^2 (*df*)	$\Delta\chi^2$	Δdf	RMSEA (CI)	Parameter	Estimates (*SE*)	Parameter (co)variances	Estimates (*SE*)
Delinquent Behavior (cont.)						Age	−0.073*** (0.018)	Age^2	0.000
						Age^2	0.0050*** (0.0011)	Intercept × Age	0.040 (0.031)
								Intercept × Age^2	−0.0076*** (0.0019)
								Age × Age^2	0.000
Externalizing	41,901.5 (2065)	42,560.3 (2074)	658.8***	9	0.10 (0.09–0.10)	Intercept	8.07*** (0.36)	Intercept	40.43*** (3.72)
						Gender	3.29*** (0.44)	Age	0.63*** (0.081)
						Age	−0.39*** (0.067)	Age^2	0.000
						Gender × Age	−0.19*** (0.039)	Intercept × Age	−1.63** (0.57)
						Age^2	0.011** (0.0038)	Intercept × Age^2	−0.083** (0.030)
								Age × Age^2	−0.012*** (0.0024)
Attention Problems	31,953.3 (2066)	32,301.3 (2074)	348.0***	8	0.08 (0.08–0.09)	Intercept	2.36*** (0.12)	Intercept	2.15*** (0.46)
						Gender	0.67*** (0.11)	Age	0.050** (0.016)
						Age	0.18*** (0.028)	Age^2	0.000
						Age^2	−0.014*** (0.0018)	Intercept × Age	0.57*** (0.087)
								Intercept × Age^2	−0.062*** (0.0060)
								Age × Age^2	0.00028 (0.00052)
Social Problems	25,843.9 (2065)	26,200.5 (2074)	356.6***	7	0.07 (0.07–0.08)	Intercept	1.23*** (0.075)	Intercept	1.26*** (0.25)
						Age	0.094*** (0.019)	Age	0.058** (0.020)
						Age^2	−0.0088*** (0.0012)	Age^2	0.000075 (0.000083)
								Intercept × Age	0.087 (0.060)
								Intercept × Age^2	−0.017*** (0.0041)
								Age × Age^2	−0.0019 (0.0013)
Thought Problems	14,535.5 (2074)			2	0.05 (0.05–0.06)	Intercept	0.26*** (0.012)	Intercept	0.15*** (0.0086)
Total Problems	53,741.1 (2065)	54,442.4 (2074)	401.3***	8	0.11 (0.10–0.12)	Intercept	21.35*** (0.68)	Intercept	164.96*** (17.49)
						Gender	4.27*** (0.98)	Age	1.66*** (0.39)
						Age	−0.38*** (0.06)	Age^2	0
						Gender × Age	−0.34** (0.09)	Intercept × Age	1.91 (2.71)
								Intercept × Age^2	−0.86*** (0.16)
								Age × Age^2	−0.010 (0.012)

Note. For all chi-squares, $N = 2{,}076$. RMSEA = root-mean-square error of approximation; CI = confidence interval.
** $p < .01$. *** $p < .0001$.

between-subjects level (Level 1) and the estimated parameter variances and covariances of the within-subject level (Level 2) as well as the values for the fit of the tested models. Except for thought problems, all normative developmental trajectories show changes across age (see Figures 1, 2, and 3), and all final models showed a significantly better fit to the data than the baseline model with only an intercept on the between-subjects and the within-subject level. The RMSEA indicates for most models an average fit, but it indicates a mediocre to poor fit for the models for Total Problems score and externalizing problems. Below, we describe the trajectories that were estimated for internalizing syndromes, externalizing syndromes, and other syndromes, respectively.

Internalizing Problems

The first entry in Table 3 describes the model for the normative developmental trajectory for the anxious/depressed syndrome. The final model, $\chi^2(2065, N = 2{,}076) = 32{,}604.3$, is significantly better than the baseline model, $\chi^2(2074, N = 2{,}076) = 32{,}875.3$, as is shown by the chi-square difference test, $\Delta\chi^2(9, N = 2{,}076) = 271.0$, $p < .0001$. The RMSEA (0.08) of the model indicates that the model fit is fair. The between-subjects level of the anxious/depressed syndrome is dependent on an intercept, gender, age, Gender × Age, and age squared effect. Figure 1 shows the normative developmental trajectory of the anxious/depressed syndrome. The initial value at age 4 is significantly different from zero (intercept = 1.73, $p < .0001$). This normative trajectory has no significant gender effect (0.26, *ns*) on the intercept. However, there is a significant gender effect on the slope (Gender × Age = −0.076, $p < .0001$), which indicates that the normative developmental trajectory is different for boys and girls. The significant linear (age) and quadratic (age squared) slope effect indicate that the normative developmental trajectory first shows an increase and thereafter shows a decrease with age (age = 0.19, $p < .0001$; age squared = −0.0093, $p < .0001$). The within-subject level is dependent on intercept, age, and age squared. These parameters account for the simple variation in the growth parameters for different individuals across the total group. The estimated (co)variances of the individual growth parameters provide us with information about the deviations of individuals' scores from the normative developmental trajectory. Of the variances of the within-subject level, only the variance of the intercept (2.16, $p < .0001$) is significant, indicating that there is variation among the individuals in the initial level of anxious and depressive problems. The significant covariances between intercept and age (Intercept × Age = 0.51, $p < .0001$) and intercept and age squared (Intercept × Age Squared = −0.055, $p < .0001$) indicate that children who start at a higher level tend to change faster than those who start at a lower level but that those with a higher initial level have a more linear decrease than those with a lower initial level.

For somatic complaints, the normative developmental trajectory is dependent on intercept, gender, age, and the interaction between gender and age. This model is significantly better than the baseline model, $\Delta\chi^2(8, N = 2{,}076) = 308.5$, $p < .0001$. The normative developmental trajectory shown in Figure 1 is significantly different from zero (intercept = 0.57, $p < .0001$) and is different for boys and girls (gender = 0.054, *ns*; Gender × Age = −0.042, $p < .0001$), and both boys and girls show an increasing trajectory over time (age = 0.054, $p < .0001$). The within-subject level is again

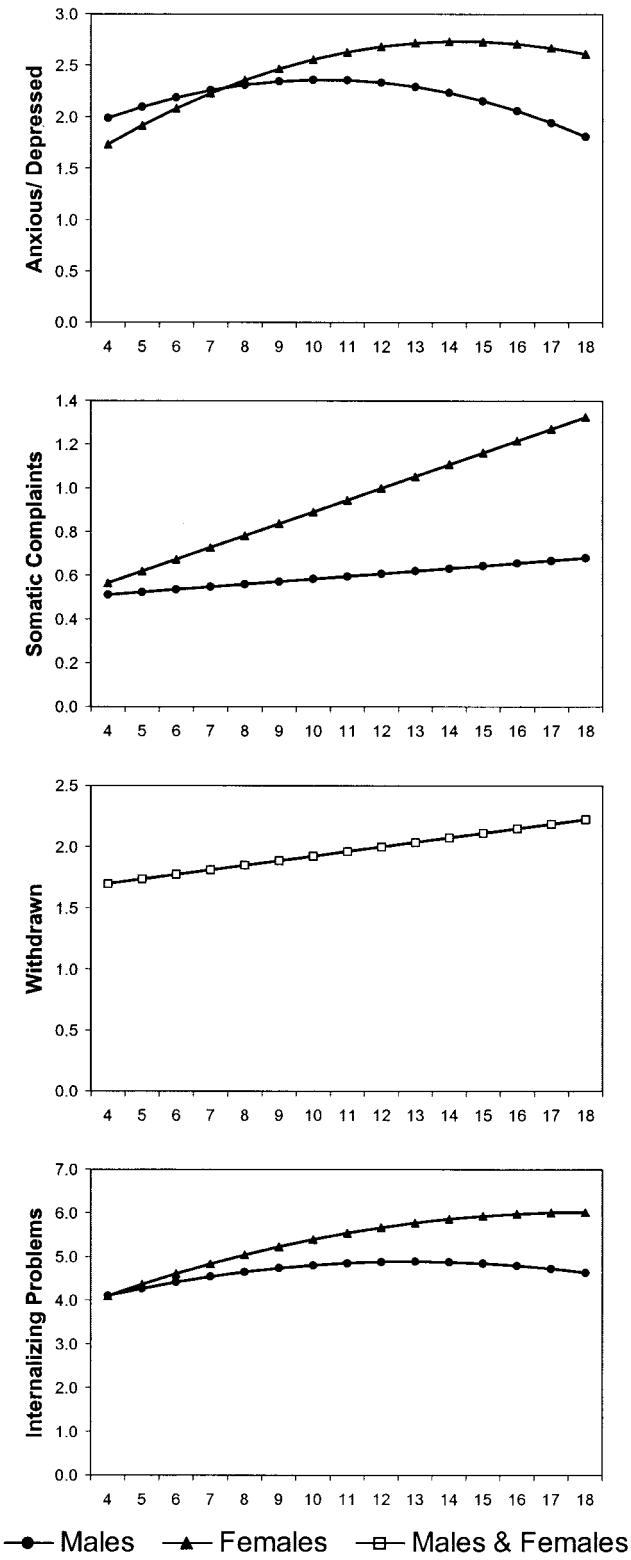

Figure 1. Normative developmental trajectories of internalizing Child Behavior Checklist syndromes and internalizing problems. Ages are shown on the *x* axis. The *y* axis represents the raw syndrome scores.

dependent on intercept, age, and age squared. Only the variances of age (0.030, $p < .0001$) and age squared (0.00024, $p < .0001$) and the covariances between intercept and age (0.057, $p < .0001$) and between age and age squared (-0.0025, $p < .0001$) are significant. This indicates that there is no significant variation in the number of somatic complaints at age 4 years. However, there is significant variation within individuals in the change of somatic complaints over age.

The final model for the withdrawn syndrome fits significantly better than does the baseline model, $\Delta\chi^2(6, N = 2{,}076) = 213.2$, $p < .05$, and is dependent on intercept and age at the between-

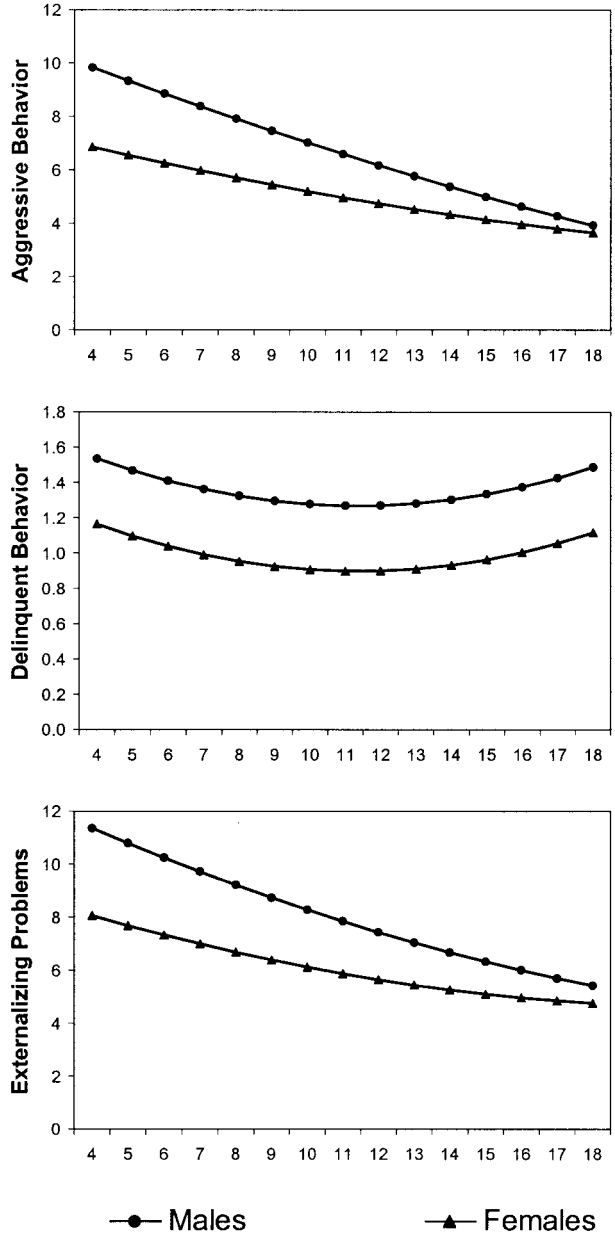

Figure 2. Normative developmental trajectories of externalizing Child Behavior Checklist syndromes and externalizing problems. Ages are shown on the *x* axis. The *y* axis represents the raw syndrome scores.

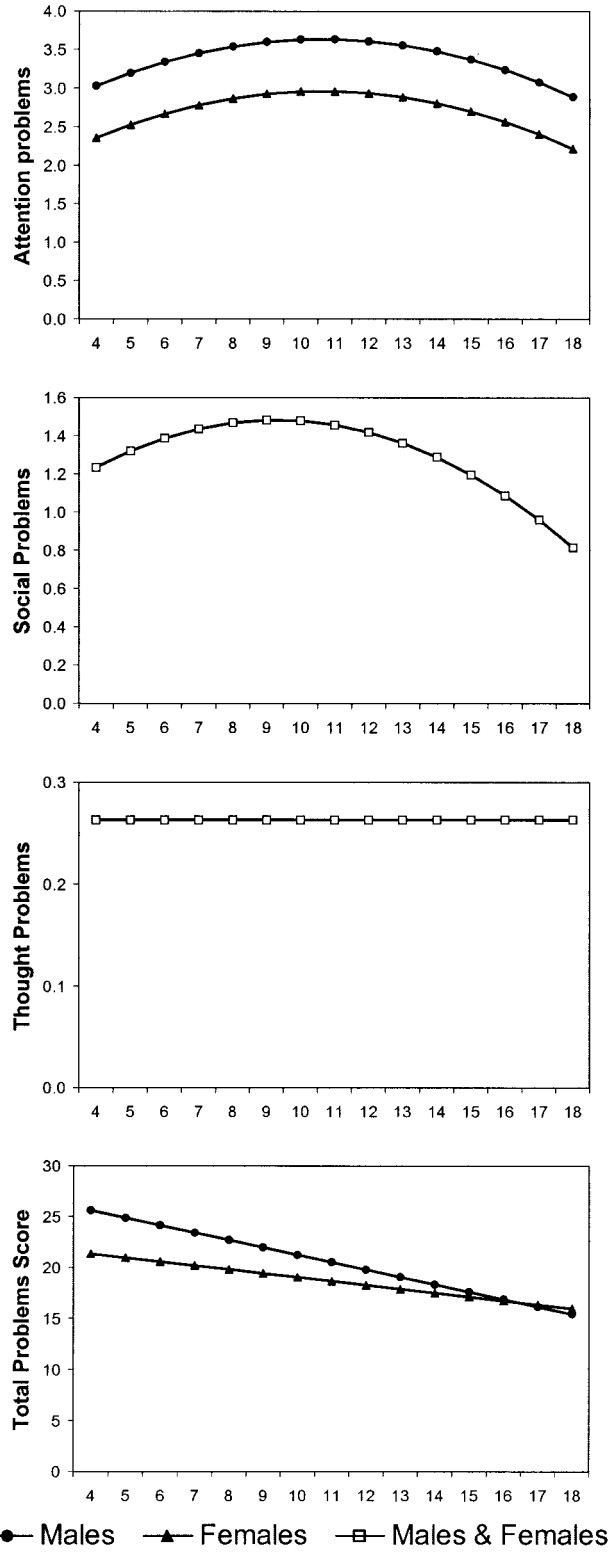

Figure 3. Normative developmental trajectories of attention problems, social problems, thought problems, and the Total Problems score. Ages are shown on the *x* axis. The *y* axis represents the raw syndrome scores.

subjects level. The normative developmental trajectory in Figure 1 shows an increase (age = 0.04, $p < .0001$) of withdrawn behavior over age that differs significantly from zero (intercept = 1.70, $p < .0001$). There is no significant gender difference in the development of withdrawn behavior. The within-subject level is again dependent on intercept, age, and age squared. The variance of the intercept (1.38, $p < .0001$) and the covariances between intercept and age (0.19, $p < .0001$) and between intercept and age squared (-0.023, $p < .0001$) are significant. This indicates that there is significant variation in the amount of problem behavior at age 4. The within-subject-level model also indicates that children who start at a higher level tend to change faster and decrease at a later measurement moment than do children who start at a lower level.

The final model for internalizing problems shows a significantly better fit than the baseline model, $\Delta\chi^2(9, N = 2{,}076) = 299.4$, $p < .0001$. The normative developmental trajectory of internalizing behavior problems is dependent on intercept, gender, age, Gender × Age, and age squared. The normative developmental trajectory (Figure 1) has no gender difference at age 4 (intercept = 4.10, $p < .0001$; gender = 0.0056, *ns*) but shows a significant effect of gender on the slope (Gender × Age = -0.099, $p < .0001$). The slope has a linear and a quadratic effect, shown in Figure 1 as a steeper increase at younger than at older ages (age = 0.27, $p < .0001$; age squared = -0.0098, $p < .01$). The within-subject level is dependent on intercept, age, and age squared. There is significant variation in the initial value of the internalizing problems (intercept = 5.17, $p < .0001$), and all the covariances are significant (Intercept × Age = 1.71, $p < .0001$; Intercept × Age Squared = -0.16, $p < .0001$; Age × Age Squared = 0.0064, $p < .0001$), indicating that children with higher initial values tend to change at a faster rate and that both children with a higher initial value and older children show a stronger decline at the end of the measurement period.

Externalizing Problems

Figure 2 shows the normative developmental trajectories of externalizing behavior problems. The normative developmental trajectory of aggressive behavior (Figure 2, Panel 1) is dependent on intercept, gender, age, Gender × Age, and age squared. The final model, $\chi^2(2065, N = 2{,}076) = 38{,}955.9$, has a significantly better fit than the baseline model, $\chi^2(2074, N = 2{,}076) = 39{,}856.0$; $\Delta\chi^2(8, N = 2{,}076) = 900.1$, $p < .0001$. Gender (2.98, $p < .0001$) has a significant effect on the intercept (6.86, $p < .0001$), which results in a different initial value for boys (initial value = 9.84) versus girls (initial value = 6.86). There is also a gender effect on the slope, which results in a difference in the development of aggressive behavior for boys and girls, with boys decreasing at a faster rate than girls (age = -0.31, $p < .0001$; Gender × Age = -0.19, $p < .0001$; age squared = 0.0059, *ns*). The within-subject level is again influenced by intercept, age, and age squared . The variances of intercept (33.13, $p < .0001$) and age (0.47, $p < .01$) are significant, which suggests that there is individual variation in the number of aggressive problems at age 4 and that there is variation in the development of the problems over age. The covariance between intercept and age (-1.72, $p < .01$) is the only covariance that is significant, which indicates that children with a lower initial value change at a faster rate than do children with a higher initial value.

The second panel of Figure 2 shows the normative developmental trajectories for delinquent behavior. These are significantly different from the baseline model, $\Delta\chi^2(8, N = 2{,}076) = 267.3$, $p < .0001$, and are dependent on intercept, gender, age, and age squared on the between-subjects level. There is a significant gender effect (0.37, $p < .0001$) on the initial value (intercept = 1.17, $p < .0001$) at age 4, resulting in a different initial value for boys (initial value = 1.54) versus girls (initial value = 1.17). There is no gender effect on the slope of the normative trajectory. Instead, the developmental trajectory of boys and girls shows the same quadratic change over time (age = -0.0073, $p < .0001$; age squared = 0.0050, $p < .0001$), resulting in a curvilinear growth for both boys and girls. The individual difference (within-subject level) is dependent on intercept, age, and age squared. The variances of the intercept (0.69, $p < .0001$) and age (0.020, $p < .01$) are significant, indicating that there are differences between individuals in the initial value and in change of delinquent behavior over time. Also, the covariance between the intercept and age squared (-0.0076, $p < .0001$) is significant, indicating that children with a lower initial score on delinquent behavior have a greater quadratic change than do children with a higher initial value on delinquent behavior.

The last panel of Figure 2 depicts the normative developmental trajectories of externalizing problems. The final model is significantly different from the baseline model, $\Delta\chi^2(9, N = 2{,}076) = 658.8$, $p < .0001$. The normative developmental trajectory is dependent on intercept, gender, age, Gender × Age, and age squared. As for all externalizing behavior problems, the initial value is significantly different for boys and girls (intercept = 8.07, $p < .0001$; gender = 3.29, $p < .0001$), which results in a different initial value of the normative developmental trajectory for boys (initial value = 11.36) versus girls (initial value = 8.07). Also, there is a significant gender effect on the slope (Gender × Age = -0.19, $p < .0001$) as well as a significant linear and quadratic effect on the slope (age = -0.39, $p < .0001$; age squared = 0.011, $p < .01$). The within-subject level indicates that individuals differ in the initial number of externalizing problems (intercept = 40.43, $p < .0001$) and that there is variation in the development of externalizing behaviors within individuals (age = 0.63, $p < .0001$). All covariances are significant (Intercept × Age = -1.63, $p < .01$; Intercept × Age Squared = -0.083, $p < .01$; Age × Age Squared = -0.012, $p < .01$), which indicates that children with lower initial values tend to change at a faster rate and also that children with lower initial levels and younger children show a stronger decrease at the end of the measurement period.

Other Problems

Figure 3 shows the normative developmental trajectories of the other syndromes (attention problems, social problems, thought problems, and Total Problems score) included in the CBCL. The first panel shows the normative developmental trajectories of attention problems. The final model of attention problems is dependent on intercept, gender, age, and age squared at the between-subjects level and has a significantly better fit than the baseline model, $\Delta\chi^2(8, N = 2{,}076) = 348.0$, $p < .0001$. The initial value of the normative developmental trajectory is different for boys (initial value = 3.03) and girls (initial value = 2.36; intercept = 2.36, $p < .0001$; gender = 0.67, $p < .0001$). There is no

gender effect on the slope, which indicates that the linear and quadratic change over time is the same for boys and girls (age = 0.18, $p < .0001$; age squared = −0.014, $p < .0001$). The within-subject level is dependent on intercept, age, and age squared, which indicates that individuals differ in the initial level of the problem behavior and in the change over time. The variances of the intercept (2.15, $p < .0001$) and age (0.050, $p < .01$) are significant. The covariances between intercept and age (0.57, $p < .0001$) and between intercept and age squared (−0.062, $p < .0001$) are also significant and indicate that children with a higher initial level change at a faster rate and that children with lower initial levels tend to show a stronger decrease at later age.

The second panel of Figure 3 shows the normative developmental trajectory of social problems, which is dependent on intercept, age, and age squared. The final model is significantly better than the baseline model, $\Delta\chi^2(7, N = 2{,}076) = 356.6$, $p < .0001$. There is no gender difference in the intercept nor in the slope. The initial value of social problems is 1.23 ($p < .0001$). After age 4 there is first a slight increase and thereafter a slight decrease (age = 0.094, $p < .0001$; age squared = −0.0088, $p < .0001$). The within-subject level of social problems is dependent on intercept, age, and age squared. The variance of intercept (1.26, $p < .0001$) and age (0.058, $p < .01$) and the covariance between intercept and age squared (−0.0117, $p < .0001$) are significant, which indicates that the subjects differ in the initial value and in the development over time.

The normative developmental model of thought problems is identical to the baseline model, $\chi^2(2074, N = 2{,}076) = 14{,}535.5$. There is no significant change in the number of thought problems over time. The only significant parameter in the between-subjects model is the intercept (0.26, $p < .0001$), which indicates that the normative developmental trajectory is significantly different from zero. Also, in the within-subject level only the variance of the intercept (0.15, $p < .0001$) is significant, indicating that subjects only differ in the initial value of the thought problems, which remains stable over time.

The last panel of Figure 3 shows the normative developmental trajectories for the Total Problems score. The final model of the Total Problems score is dependent on intercept, gender, age, and Gender × Age on the between-subjects level and on intercept, age, and age squared on the within-subject level. This final model has a significantly better fit than the baseline model, $\Delta\chi^2(8, N = 2{,}076) = 401.3$, $p < .0001$, and is significantly different for boys and girls. There is a significant gender effect (4.27, $p < .0001$) on the intercept (21.35, $p < .0001$) and on the slope (age = −0.38, $p < .0001$; Gender × Age = −0.34, $p < .0001$). The within-subject-level variances and covariances indicate that individuals differ in the initial value and in the amount of change over time; the variances of the intercept (164.94, $p < .0001$) and age (1.66, $p < .0001$) as well as the covariance between intercept and age squared (−0.86, $p < .0001$) are significant.

Discussion

The aim of this study was to describe the normative developmental trajectories of behavioral and emotional problems during childhood and adolescence. This study accounts for most of the problems typically associated with this type of research. The trajectories were described in a large longitudinal sample including multiple cohorts of boys and girls aged 4 to 18 years. Using Time 1 SES and CBCL Total Problems scores, we found significant differences in SES between the dropout groups but no indication for selective attrition for the Total Problems score. These findings suggest that children and youths with lower SES are somewhat underrepresented in the longitudinal sample but that these children do not show more problem behavior. To further analyze possible dropout effects, we also tested whether the normative developmental trajectories of the remainders and the random dropouts differed from the normative developmental trajectories of the total sample. This test indicated that the parameters of the developmental trajectories of the different dropout groups were in each other's range of confidence interval, which suggests that the deviation from the normative development is not significant for different dropout groups in the sample. Finally, we added a variable to the multilevel growth curve model that represented the subjects' dropout group. This variable was not significant in any of the normative developmental trajectories. In sum, the net effect of the selective dropout of a group with lower SES is visible neither in initial problem scores nor in the normative developmental trajectories. The design of the study corrected for possible cohort and period effects, because all ages except ages 4 and 5 were measured multiple times in different cohorts and at different time points. The analytic method used also corrects for the dependencies among observations created by repeated measurements.

This study gives new insights in the development of behavioral and emotional problems of children and adolescents in the general population. Some of our findings are in stark contrast to our expectations, whereas other expectations were confirmed.

Internalizing Problems

The normative developmental trajectory for the CBCL Internalizing scale showed a curvilinear increase for both girls and boys over time. In childhood, the number of internalizing problems did not differ between boys and girls, whereas in adolescence the developmental trajectory for girls showed a higher average level than that for boys. The same trajectories for the CBCL Internalizing scale were found in other longitudinal studies (Keiley et al., 2000; Stanger & Verhulst, 1995). These findings reflect that internalizing problems are the same for boys and girls in childhood but have a different developmental trajectory thereafter. The different developmental trajectories for boys and girls are likely to be explained by differences in pubertal development or different coping styles for boys versus girls (e.g., Zahn-Waxler, Klimes-Dougan, & Slattery, 2000).

Confirming our expectations, we found a developmental trajectory for the Withdrawn scale reflecting increasing withdrawal with increasing age. On the basis of small differences between boys and girls found in cross-sectional studies (Achenbach, 1991; Verhulst et al., 1996), we also expected a gender difference, with girls showing more withdrawn behavior than boys. However, this could not be confirmed in the present study. The trajectory found for the Withdrawn scale confirms earlier observations that during the transition into adolescence children spend increasing amounts of time alone or with friends, show a dramatic drop in time they spend with their parents (Alsaker, 1996), and disclose less of their inner world to their parents.

As expected, we found different trajectories for boys and girls for the Somatic Complaints scale. Contrary to our expectation, we found that Somatic Complaints scores increased with increasing age for girls, whereas the trajectory for boys was less influenced by age. Scores on the Somatic Complaints scale for girls were twice as high in adolescence than in childhood. In a cross-sectional study, Eminson, Benjamin, Shortall, and Woods (1996) also found significantly more physical symptoms in girls versus boys and more symptoms in older versus younger girls, but Egger et al. (1999) and Taylor et al. (1996) did not. An increase such as the one found in the present study and in the study by Eminson et al. (1996) may have been missed in the Egger et al. (1999) and Taylor et al. (1996) studies because of the limited age ranges covered by these studies (ages 9–16 years and 12–16 years in the Egger et al., 1999 and Taylor et al., 1996, study, respectively) as well as the fact that these studies looked at single items instead of scales composed of multiple items, as the present study did. The increase of somatic complaints with age for girls may be a precursor of the often reported gender differences in somatic complaints between men and women in adulthood (e.g., Rief, Hessel, & Braehler, 2001). The present study's results suggest that the gender difference in somatic complaints in adulthood may have originated in childhood.

For the Anxious/Depressed scale, we found a normative developmental trajectory for boys that showed first a slight increase and thereafter a decrease. The trajectory for boys showed in childhood and in adolescence nearly the same level of problems, which confirmed our expectation that anxious–depressive problems are fairly stable for boys. For girls, the normative developmental trajectory showed a higher level of problems than for boys and the expected increase with increasing age (Angold et al., 1998). However, after an initial increase, the trajectory slightly decreased. Informant effects may explain the discrepancies between our expectations and our findings. Many symptoms of adolescent anxiety and depression are covert, and parents may be less accurate informants of these problems than are the adolescents themselves (Compas et al., 1997; Mesman & Koot, 2000a, b; Verhulst & van der Ende, 1992). For example, in a large national sample of referred and nonreferred youths, Compas et al. (1997) found that parents reported for the nonreferred youths fewer problems on the anxious/depressed syndrome with increasing age. The nonreferred youths' self-reported scores on the Anxious/Depressed scale showed a gradual increase with increasing age, and girls reported more problems than boys. This suggests that the normative developmental trajectory of anxiety and depression reported by parents may be different from the normative developmental trajectory reported by youths.

Externalizing Problems

As expected, boys and girls differed in their trajectories for Externalizing scores. Boys and girls both showed normative developmental trajectories of scores that decreased with age, with boys showing more problems than girls, throughout the whole measurement period. Similar declining trajectories were found by Stanger and Verhulst (1995), Keenan and Shaw (1997), and Silverthorn and Frick (1999) but not by Keiley et al. (2000), who found no gender differences in the declining externalizing trajectory.

As expected, the normative developmental trajectory for the Aggressive Behavior scale decreases with age for both boys and girls. In childhood, boys had higher levels of parent-reported aggressive behaviors than did girls. However, aggressive behaviors decreased at a much faster rate with age in boys than in girls, with nearly no gender difference left at 18 years. This finding corresponds with findings of many earlier studies that reported higher levels of aggression in boys versus girls in childhood (Crick & Dodge, 1996; Keenan & Shaw, 1997). Contradicting the popular belief that as boys become older they increase the frequency of their oppositional and physically aggressive behavior, the present study shows a decline in aggressive behavior over time even more so in boys than in girls. According to our study, aggressive behaviors shown by children and adolescents are relatively transitory in nature and are resolved by the beginning of adulthood.

We found a normative developmental trajectory with a curvilinear increase for the Delinquent Behavior scale for both boys and girls, with boys showing more rule breaking behaviors than girls. On the basis of a study by Moffitt et al. (1996) on antisocial and delinquent acts, we expected an increase in the developmental trajectory for delinquent behavior for boys in adolescence. According to Moffitt et al. (1996), almost 30% of the boys commit antisocial or delinquent acts only in adolescence, whereas 5% of the boys commit antisocial or delinquent acts throughout life. The present study's results confirm the presence of a relatively large group of boys as well as girls committing delinquent acts in adolescence. However, the normative developmental trajectory of delinquent behaviors cannot give conclusive evidence for the existence or nonexistence of a life-course-persistent group that commits antisocial or delinquent acts starting at a young age and lasting throughout life.

For attention problems, we found a trajectory for both boys and girls that increased until age 11 and declined thereafter. Although boys and girls followed the same trajectory, boys showed more attention problems than girls. Our findings may well be explained by current theory on the development of attention problems and overactivity, with increases in these problems as the demands on impulse control and response inhibition increase during childhood, and with a decrease in these problems with the growing development of self-regulation in adolescence (Barkley, 1997; Hart et al., 1995).

Implications and Limitations

The trajectories as described in this study reflect the normative development expected if child and adolescent problems are repeatedly assessed using parent reports. The trajectories found in this study provide a basis against which deviations from the expected developmental course can be identified.

Some of the trajectories found in the present study throw a new light on the development of problem behaviors across childhood and adolescence, as illustrated by the decrease in scores on the Aggressive Behavior scale and the increase in scores on the Withdrawn and Somatic Complaints scales. We know of no other studies demonstrating the curvilinear course of attention problems.

The present study is not without limitations. A main limitation is the generalizability of findings. The longitudinal design resulted in some selective dropout of study participants. The SES of the dropouts was slightly, though significantly, lower than the SES of

the remainders, indicating that the remainders had slightly better socioeconomic background than the dropouts. However, post hoc analyses indicated that it is unlikely that this dropout affected the estimated developmental trajectories to any significant extent.

The study population was a random sample of mainly Caucasian children and adolescents living in the Netherlands. Crijnen, Achenbach and Verhulst (1999) compared CBCL scores for 11 different cultures and concluded that cultural effects were minimal and that the CBCL can be used to obtain reports of problem behaviors that can be compared across countries.

The design of the study created the opportunity to correct for cohort and period effects. A disadvantage of this design is that fewer measurement points were available for the individuals in the older birth cohorts and at the younger age (4–5 years). Notwithstanding these limitations in the design, the analyses have a considerable power (Hedeker, Gibbons, & Waternaux, 1999).

Another limitation of this study is the reliance on only parental reports to assess psychopathology. Parents may be unaware of their child's rule-breaking behavior, especially as their child becomes an adolescent. The same may be true for internalizing problems. Parents are not the most reliable source to report anxious and depressed feelings of their children (Mesman & Koot, 2000a, b; Verhulst & Van der Ende, 1992). Therefore, replications are essential to assess the generalizability of the present findings to other informants, such as teachers and youths themselves.

The developmental trajectories of problem behavior obtained in this study may constitute the background for additional and smaller range studies. The development of problem behavior in specific clinical or risk groups may be contrasted with the present normative data, enabling the detection of groups of deviating individuals. In addition, it will be interesting to identify children within the present sample who follow individual developmental trajectories that deviate from the normative trajectories found in this study (e.g., Nagin & Tremblay, 1999).

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 Child Profile*. Burlington: University of Vermont, Department of Psychiatry.

Alsaker, F. D. (1996). Annotation: The impact of puberty. *Journal of Child Psychology and Psychiatry, 37,* 249–258.

Angold, A., Costello, E. J., & Worthman, C. M. (1998). Puberty and depression: The roles of age, pubertal status and pubertal timing. *Psychological Medicine, 28,* 51–61.

Angold, A., & Rutter, M. (1992). Effects of age and pubertal status on depression in a large clinical sample. *Development and Psychopathology, 4,* 5–28.

Barkley, R. A. (1997). Attention-deficit/hyperactivity disorder, self-regulation, and time: Toward a more comprehensive theory. *Journal of Developmental and Behavioral Pediatrics, 18,* 271–279.

Bell-Dolan, D. J., Last, C. G., & Strauss, C. C. (1990). Symptoms of anxiety disorders in normal children. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 759–765.

Bernstein, G. A., Borchardt, C. M., & Perwien, A. R. (1996). Anxiety disorders in children and adolescents: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1110–1119.

Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dahl, R. E., et al. (1996). Childhood and adolescent depression: A review of the past 10 years, Part I. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1427–1439.

Boyle, M. H., & Willms, J. D. (2001). Multilevel modelling of hierarchical data in developmental studies. *Journal of Child Psychology and Psychiatry, 42,* 141–162.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Cairns, R. B., & Cairns, B. D. (1984). Predicting aggressive patterns in girls and boys: A developmental study. *Aggressive Behavior, 10,* 227–242.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Gariépy, J. L. (1989). Growth and aggression: I. Childhood to early adolescence. *Developmental Psychology, 25,* 320–330.

Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry, 36,* 113–149.

Cantwell, D. P. (1996). Attention deficit disorder: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 978–987.

Cicchetti, D., & Toth, S. L. (1998). The development of depression in children and adolescents. *American Psychologist, 53,* 221–241.

Cohen, P., Cohen, J., Kasen, S., Velez, C. N., Hartmark, C., Johnson, J., et al. (1993). An epidemiological study of disorders in late childhood and adolescence: I. Age- and gender-specific prevalence. *Journal of Child Psychology and Psychiatry, 34,* 851–867.

Compas, B. E., Oppedisano, G., Connor, J. K., Gerhardt, C. A., Hinden, B. R., Achenbach, T. M., & Hammen, C. (1997). Gender differences in depressive symptoms in adolescence: Comparison of national samples of clinically referred and nonreferred youths. *Journal of Consulting and Clinical Psychology, 65,* 617–626.

Craske, M. G. (1997). Fear and anxiety in children and adolescents. *Bulletin of the Menninger Clinic, 61,* A4–A36.

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development, 67,* 993–1002.

Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1999). Problems reported by parents of children in multiple cultures: The Child Behavior Checklist syndrome constructs. *American Journal of Psychiatry, 156,* 569–574.

De Groot, A., Koot, H. M., & Verhulst, F. C. (1994). Cross-cultural generalizability of the Child Behavior Checklist cross-informant syndromes. *Psychological Assessment, 6,* 225–230.

Egger, H. L., Costello, E. J., Erkanli, A., & Angold, A. (1999). Somatic complaints and psychopathology in children and adolescents: Stomach aches, musculoskeletal pains, and headaches. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 852–860.

Eminson, M., Benjamin, S., Shortall, A., & Woods, T. (1996). Physical symptoms and illness attitudes in adolescents: An epidemiological study. *Journal of Child Psychology and Psychiatry, 37,* 519–528.

Farrington, D. P. (1991). Longitudinal research strategies: Advantages, problems, and prospects. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 369–374.

Fergusson, D. M. (1998). Stability and change in externalising behaviours. *European Archives of Psychiatry and Clinical Neuroscience, 248,* 4–13.

Fleming, J. E., & Offord, D. R. (1990). Epidemiology of childhood depressive disorders: A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 571–580.

Frick, P. J., Lahey, B. B., Loeber, R., Tannenbaum, L., Van Horn, Y., Christ, M. A. G., et al. (1993). Oppositional defiant disorder and conduct disorder: A meta-analytic review of factor analyses and cross-validation in a clinic sample. *Clinical Psychology Review, 13,* 319–340.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1036–1045.

Gomez, R., Harvey, J., Quick, C., Scharer, I., & Harris, G. (1999). *DSM–IV* AD/HD: Confirmatory factor models, prevalence, and gender and age

differences based on parent and teacher ratings of Australian primary school children. *Journal of Child Psychology and Psychiatry, 40*, 265–274.

Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B., Green, S. M., & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology, 23*, 729–749.

Hedeker, D., Gibbons, R. D., & Waternaux, C. (1999). Sample size estimation for longitudinal designs with attrition: Comparing time-related contrasts between two groups. *Journal of Educational and Behavioral Statistics, 24*, 70–93.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin, 121*, 95–113.

Keiley, M. K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A cross-domain growth analysis: Externalizing and internalizing behaviors during 8 years of childhood. *Journal of Abnormal Child Psychology, 28*, 161–179.

Kraemer, H. C., Yesavage, J. A., Taylor, J. L., & Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry, 157*, 163–171.

Laitinen-Krispijn, S., Van der Ende, J., & Verhulst, F. C. (1999). The role of the pubertal progress in the development of depression in early adolescence. *Journal of Affective Disorders, 54*, 211–215.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS System for mixed models*. Cary, NC: SAS Institute.

Loeber, R., & Farrington, D. P. (1995). Longitudinal approaches in epidemiological research of conduct problems. In F. C. Verhulst & H. M. Koot (Eds.), *The epidemiology of child and adolescent psychopathology* (pp. 309–336). Oxford, England: Oxford University Press.

Loeber, R., Wung, P., Keenan, K., Giroux, B., Stouthamer-Loeber, M., Welmoet, B., et al. (1993). Developmental pathways in disruptive child behavior. *Development and Psychopathology, 5*, 103–133.

Longford, N. T. (1993). *Random coefficient models*. Oxford, England: Clarendon Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

Mesman, J., & Koot, H. M. (2000a). Child-reported depression and anxiety in preadolescence: I. Associations with parent- and teacher-reported problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1371–1378.

Mesman, J., & Koot, H. M. (2000b). Child-reported depression and anxiety in preadolescence: II. Preschool predictors. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1379–1386.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674–701.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology, 13*, 355–375.

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology, 8*, 399–424.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development, 70*, 1181–1196.

Raudenbush, S. W., & Chan, W. (1992). Growth curve analysis in accelerated longitudinal designs. *Journal of Research in Crime and Delinquency, 29*, 387–411.

Rief, W., Hessel, A., & Braehler, E. (2001). Somatization symptoms and hypochondriacal features in the general population. *Psychosomatic Medicine, 63*, 595–602.

Silverthorn, P., & Frick, P. J. (1999). Developmental pathways to antisocial behavior: The delayed-onset pathway in girls. *Development and Psychopathology, 11*, 101–126.

Stanger, C., Achenbach, T. M., & Verhulst, F. C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Development and Psychopathology, 9*, 43–58.

Stanger, C., & Verhulst, F. C. (1995). Accelerated longitudinal designs. In F. C. Verhulst & H. M. Koot (Eds.), *The epidemiology of child and adolescent psychopathology* (pp. 385–405). Oxford, England: Oxford University Press.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling, 5*, 411–419.

Taylor, D. C., Szatmari, P., Boyle, M. H., & Offord, D. R. (1996). Somatization and the vocabulary of everyday bodily experiences and concerns: A community study of adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 491–499.

Tremblay, R. E. (2000). The development of aggressive behaviour during childhood: What have we learned in the past century? *International Journal of Behavioral Development, 24*, 129–141.

Van Westerlaak, J. H., Kropman, J. A., & Collaris, J. W. M. (1975). *Beroepenklapper* [Manual for occupational level]. Nijmegen, the Netherlands: Instituut voor Sociologie.

Verhulst, F. C., Akkerhuis, G. W., & Althaus, M. (1985). Mental health in Dutch children: I. A cross-cultural comparison. *Acta Psychiatrica Scandinavica, 323*, 1–108.

Verhulst, F. C., Koot, H. M., & Berden, G. F. M. G. (1990). Four-year follow-up of an epidemiological sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 440–448.

Verhulst, F. C., & van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry, 33*, 1011–1023.

Verhulst, F. C., van der Ende, J., & Koot, H. M. (1996). *Handleiding voor de CBCL/4–18* [Manual for the CBCL/4–18]. Rotterdam, the Netherlands: Erasmus University and Department of Child and Adolescent Psychiatry, Sophia Children's Hospital.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology, 10*, 395–426.

Zahn-Waxler, C., Klimes-Dougan, B., & Slattery, M. J. (2000). Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and Psychopathology, 12*, 443–466.

Received March 18, 2002
Revision received August 25, 2002
Accepted September 15, 2002 ■